

The origins of de Finetti's critique of countable additivity

Eugenio Regazzini*

University of Pavia

Abstract: Bruno de Finetti was one of the most convinced advocates of finitely additive probabilities. The present work describes the intellectual process that led him to support that stance and provides a detailed account both of the first paper by de Finetti on the subject and of the ensuing correspondence with Maurice Fréchet. Moreover, the analysis is supplemented by a useful picture of de Finetti's interactions with the international scientific community at that time, when he elaborated his subjectivistic conception of probability.

1. Introduction

Finitely additive probabilities are indissolubly linked to the name of Bruno de Finetti (1906-1985). Indeed, he has been one of the most convinced advocates of finite additivity, since he has started working on the mathematical formulation that he proposed, in 1930, for his subjectivistic conception of probability. Most of the recent contributions to this topic in the literature rely on (English translations of) late works by de Finetti, instead of considering his early papers containing a wealth of fresh and original ideas. A typical feature of de Finetti's late works is that they generally aim at providing critical syntheses of his original way of thinking about crucial problems concerning: Probability, Induction, Statistics, Insurance, Economics, Politics, to say nothing of the philosophical debate at large. Hence, in these writings he makes limited use of mathematical formalism, omits precise references of an historical nature, whilst he often jumps at the chance of both controversial amusing hints and sharp provocative cues. Nevertheless, late works share with the early ones the feature of being faultless from the point of view of logical and conceptual accuracy. It is not, then, surprising that, as far as finite additivity is concerned, essays on de Finetti's work often reduce the topic either to an intellectual activity or, at best, to an issue of a mere mathematical taste. In the latter case, special attention is given to connections and consequences of a formal nature. So, they generally fail to

Dipartimento di Matematica "F. Casorati" Università di Pavia (Italy) e-mail: eugenio.regazzini@unipv.it

*Also affiliated to CNR-IMATI, Milano (Italy). This work is partially supported by the Italian Ministry of University and Research, grant n. 2008MK3AFZ. The author also wishes to thank Antonio Lijoi and an anonymous Referee for their comments that led to an improved exposition.

AMS 2000 subject classifications: 60-03, 60A05

Keywords and phrases: Finitely additive probabilities; Processes with independent and stationary increments; Subjective probabilities

shed some light on the intellectual efforts that led him to share that seemingly singular position. In fact, the authentic motivations supporting de Finetti's stance in probability can be found in his early works. These, among other things, are tersely written and contain accurate formulations of theorems thus allowing one a more sound understanding of his innovations.

The main purpose of the present work is to discuss the first paper that, at the best of this author's knowledge, de Finetti devoted to the analysis of the effects of considering as admissible finitely additive laws. See [14]. The title of the paper, "*Sui passaggi al limite nel calcolo delle probabilità*", could be translated into "*On the limit processes in the calculus of probability*" and evokes the continuity property of countably additive laws. Its content is carefully described and critically annotated in Section 4. This follows a discussion, in Section 2, on the development of de Finetti's ideas and achievements in probability theory between 1927, the year of his graduation, and 1930, the year of the publication of "*Sui passaggi ...*". For a more comprehensive illustration of de Finetti's work, see [9]. Section 3 provides a concise description of the mathematical theory of probability deduced from de Finetti's *coherence principle*. See [22].

The present author is not aware of how "*Sui passaggi ...*" was received by the Italian scientific community. However, the fact that it was not published in the journal *Rendiconti della Reale Accademia dei Lincei* might conceal a cold reaction by the most influential Italian probabilist of the time, namely Francesco Paolo Cantelli (1875-1966). It is well-known that Cantelli was not enthusiastic about non countably additive probabilities. See [4] and [5, 6, 7]. A public reaction on the spur of the moment came from the famous French mathematician Maurice Fréchet (1878-1973) who initiated an interesting correspondence with de Finetti gathered into four published notes. See [15, 16], [29, 30]. The reading of this correspondence provides a vivid insight into the stances of the two "competitors" that reflect two different ways of thinking still of great interest. Then, this correspondence is reported and annotated in the present paper as well. See Section 5.

I am delighted I am given the chance to write this work in honour of Joe Eaton who devoted a distinguished part of his scientific research to the foundations of probability and statistics.

2. De Finetti and other probabilists at the end of the twenties

De Finetti started coping with the fundamental task of formulating a satisfactory and general theory of probability right after his graduation. Almost in the same period he initiated his studies on the sequences of exchangeable events and on functions with independent, stationary increments (f.i.s.i., for short). But, whilst his investigations on these specific topics proceeded rapidly – cf., e.g., [49] and [1] – the research on the foundations of probability appeared somewhat bristling with difficulties. The major challenge arose from the mathematical formulation

of the subjectivistic conception of probability as expounded, from a philosophical viewpoint, in [21], an essay that had already been drawn up at the beginning of 1930. That time lag did not restrain him from reaching considerable achievements, but led him to a retrospective critical reflection on some of his own results, that he had obtained within the realm of countably additive probabilities. This happened during the second half of 1929, when he reached the conclusion that such a condition is not necessary in order that a function on a class of events can be viewed as a probability. Indeed the presentations, at the *R. Accademia dei Lincei*, of [12] and [13] are the 7th of October and the 1st of December, respectively. The precise deduction of the characterizing properties of a probability law, from the subjectivistic standpoint, were announced in both the replies to Fréchet. See [15, 16] and [22] for the final version.

It is worth providing some further insight into this story. By assuming the continuity of the law of a f.i.s.i. based on Gaussian finite-dimensional distributions, in [11] de Finetti had stated that almost every trajectory turns out to be nowhere differentiable: a result today ascribed to Paley, Wiener and Zygmund. See [45]. Subsequently, he had shown that the probability distribution (p.d.) of X_t is continuous, for each $t > 0$, whenever $(X_t)_{t \geq 0}$ is a continuous f.i.s.i.. See [12]. A first clear statement of his concerns about the suitability of an analysis confined to considering random functions governed by countably additive laws can be found in [13] where he deals with the problem of determining the p.d. of the integral, on $[0, t]$, of a f.i.s.i.. See next Section 4. It is in this last essay that he announces the preparation of *Sui Passaggi*, with the purpose of investigating the issue from a more general viewpoint. It seems fair to say that *Sui Passaggi* marks a turning point in de Finetti's mathematical treatment of probability. After hinting at the possible effects of such an afterthought on his own work, we linger on describing the boundary conditions within which it matured.

As to the Italian scientific community, it has been already mentioned that Cantelli claimed to be unconditionally in favour of countable additivity. As an example, in his celebrated paper on the strong law of large numbers – see [4] – he had declared “Such an assumption plainly cannot raise objections from a theoretical viewpoint, and corresponds to the feeling that probability, viewed from an empirical angle, arouses in us”. He made no mention of the fundamental problem of the existence of a countably additive extension of sequences of assigned finite-dimensional laws of (X_1, \dots, X_n) , for $n = 1, 2, \dots$, to obtain a probability law for $(X_n)_{n \geq 1}$. See also [50] and [48]. We do not know whether he ignored or not the existence of such a problem that was completely solved later by Kolmogorov. See [40]. In fact, it was only in 1932 that Cantelli proposed a measure-theoretic approach with an explicit view to proving existence of extensions like the previous one. See, in particular, [44], which includes one of the main achievements within the so-called *Cantelli abstract theory of probability*. As far as the issue of interpreting the meaning of probability is concerned, Cantelli swung between an empirical interpretation and a bent for find-

ing conceptual connections between his abstract theory and the classical Laplace definition. It is for sure that he had no sympathy for subjectivistic interpretations.

As regards his interactions with the broader international scientific community, de Finetti has been keeping up correspondences with some of the foreign foremost scholars such as Fréchet, Lévy, Khinchin and Kolmogorov since the end of the 20's. It is well-known, for example, that Alexandr Y. Khinchin (1894–1959) took interest in the study of sequences of exchangeable events during and after the *8th International Congress of Mathematicians* held in Bologna, 3–10 September, 1928. See [34, 35]. Andrei N. Kolmogorov (1903–1987), following the theory devised by de Finetti in '29, obtained the renowned representation of the characteristic function of X_1 , when $(X_t)_{t \geq 0}$ is a f.i.s.i. and X_1 has finite variance. The starting point for this research is in [17], whereas the Kolmogorov contribution is contained in [38, 39]. Also the generalization due to Khinchin of the Kolmogorov statement can be viewed, unlike that given by Lévy, as a by-product of the de Finetti approach to functions with independent increments. See [36] and [42]. In the period 1929–1930, de Finetti focused on a paper by Kolmogorov concerning the representation of associative means. See [37] and [23]. Both Fréchet and Evgeny E. Slutsky (1880–1948) were mentioned by de Finetti for contributions to stochastic convergence: the former for the study of convergence of random elements in abstract spaces (see [15]), the latter for the use of the term *stochastic* that de Finetti adopted to designate convergence in probability (see [14]). As to Paul Lévy (1886–1971), it is well-known that he conceived and developed his fundamental contribution to f.i.s.i.'s independently of both de Finetti and Kolmogorov. See [43]. On de Finetti's side, Lévy's *Calcul des Probabilités* (1925) and Castelnuovo's *Calcolo delle Probabilità e Applicazioni* (1919) were the sole existing reference books at the end of the twenties, gathering the essentials of the theory and calculus of probability in a systematic way. Like Cantelli, Lévy was inclined to restrict probability laws to countably additive functions on events, and tried to justify his position in the final part of the book.

As to the meaning and the interpretation of the concept of probability, the span of time we are dealing with overlaps with the success of the so-called empirical conceptions, according to which probability is related to frequency. This position was defended, in those years, by great scientists like Guido Castelnuovo (1865–1952) and Richard von Mises (1883–1953). See, for example, [53]. De Finetti had found the empirical arguments incomplete and inadequate since the very beginning of his own approaching the probabilistic studies. Unsatisfied, he followed a different path that led him to the formulation of a radically subjectivistic theory, with the consequence that countable additivity is not necessary in order that a set function may be considered a probability. A brief description of de Finetti's theory will be sketched in the next section. We conclude the present one with a digression about the spread of basic elementary concepts in the literature of the day, by means of an example. The research into the subject has been suggested by the reading of *Sui Passaggi* and regards, in particular, definition and properties of convergence in

probability.

In Section 8 of that paper, de Finetti reminds the reader of a definition given in [13] – the definition of *convergence in probability* of sequences of random numbers – and proves that convergence in probability entails *convergence in distribution*. It will be explained, in Section 4, why he had been led to deal with this arrangement of topics. Here we want to comment on the fact that in neither of the two papers does he mention any reference and he proposes to designate the concept by the locution *convergenza stocasticamente uniforme* (stochastically uniform convergence). Combining this circumstance with the fact that de Finetti was used to carefully mention references, including those having little bearing on the development of his own research, one is led to conclude that the concept of convergence in probability was not yet a part of the probabilistic literature at the time. This may seem amazing. It is worth recalling that the topic of convergence of sequences of random numbers had generated heated controversies, not yet dulled at the end of the twenties. A significant part of the contention can be traced back to Cantelli’s determination in claiming his priority, over Emile Borel (1871–1956), for the formulation of the strong law of large numbers. See, for example, [50] and references therein. Stimulated by the draft of the present paper, we have thoroughly examined the literature of that time and we have got to the conclusion that de Finetti was probably the first to consider sequences having a random number as a limit. Moreover, one should acknowledge his priority in proving that convergence in probability is stronger than convergence in distribution. The statement on page 25 of [50], according to which “Cantelli (1916) had anticipated Slutsky (1925) in introducing a random variable (rather than just a constant) as a limit ... of convergence in probability”, does not seem correct. Indeed, on the one hand, Cantelli in [3] deals with constant limits only – see also [5] – and, on the other hand, Slutsky in [51] considers only sequences of the type $(X_n - m_n)_{n \geq 1}$ converging to zero for some sequence $(m_n)_{n \geq 1}$ of real numbers (*stochastische Asymptote*). A more delicate analysis must be devoted to the Fréchet work on convergence of random sequences. It is Fréchet himself who gave notice, in his comments on *Sui Passaggi* that were presented on the 3rd of July 1930 at the *Reale Istituto Lombardo*, of a work of his having some points of contact with de Finetti’s paper. He did not mention the title of his article, which was referred to as a summary of his recent courses, and announced it was going to appear in the (Italian) statistics journal *Metron*. Indeed, it was published in the last issue of that year. See [29] and [31]. In the meantime, de Finetti has come by the Fréchet paper: this is witnessed by the fact he touches upon it in his reply to the Fréchet first group of comments on *Sui Passaggi*. De Finetti confines himself to saying that, apart from what is consequence of the assumption of countable additivity, he has particularly appreciated the study of the convergence of sequences of random elements taking values in abstract metric spaces. See the last part of [31]. The first sections contain accurate descriptions of the concepts of convergence in probability and of almost sure convergence for sequences of random numbers. As to

the former, the definition is obviously the very same as that previously proposed, under different name, by de Finetti. Moreover, also the Fréchet work contains a proof of the fact that convergence in probability entails convergence in distribution. The funny thing is that he omits giving precise references to the points of contact with *Sui Passaggi*. In fact, he just mentions works by Slutsky, Dell'Agnola and de Finetti, without specifying either titles or other bibliographical data. He justifies these omissions by writing that “... nous n'avons pu consulter assez librement et complètement les publications de M. Slutsky ... Par contre les mémoires de M.M. Dell'Agnola et de Finetti ne nous paraissent pas consacrés aux mêmes questions que celles qui ont été traitées ici.” Such a comment is not appropriate either for the Dell'Agnola paper [27] or for *Sui Passaggi*. But, while the former will be mentioned in Chapter 5 of [32], de Finetti's contribution will be ignored even on that occasion, in spite of the anything but negligible overlapping. So, de Finetti was probably the first to introduce a general definition of convergence in probability and to study its connections with convergence in law, but Fréchet was closely following him on this path. Situations of this type are not infrequent in de Finetti's scientific production. Other noteworthy examples are represented by: (i) the concept of infinitely divisible law, introduced to encapsulate the characterizing feature of the law of X_1 in a stochastic f.i.s.i. $(X_t)_{t \geq 0}$ (see [11, 17]); (ii) the completion of the continuity theorem for characteristic functions, to provide the first proof of the renowned representation theorem for exchangeable events (see [19]); (iii) the extension of the Kolmogorov theorem for associative means (see [23]); (iv) the anticipation of the so-called Kendall's τ coefficient, in a general study on correlation and concordance (see [24]); (v) a pioneering use, almost contemporaneously with Lévy, of the notions of martingale, stopping time and optional sampling, to reformulate the Lundberg–Cramér theory of risk, without coining any new locution to designate them (see [25]). The recurrence of these circumstances is the result of a combination of his special mathematical inventiveness with the fact that, as he wrote himself, “he was interested in mathematics meant as a tool for applications ... and for the investigation of conceptual and critical issues ..., rather than as formalism or as an abstract subject, axiomatized and withdrawn in itself.” See Page XVIII of [26]. This attitude prevented him from isolating results that were merely proved with a view to the solution of a more general problem. It also made it difficult to acknowledge the paternity of a number of scientific innovations.

3. De Finetti's coherence principle

Responding to Fréchet's second series of remarks, de Finetti claims that the most significant point of the question on additivity lies in the need of proofs for the properties (of probability) that one wishes to affirm. See [16]. For those who, like us, are accustomed to affirm those properties by means of axioms, de Finetti's recommendation might appear as the fruit of an outdated way of thinking. However, it is plain to see that it represents the most reasonable way to settle the dispute

constructively. Then, he asks every author to provide a proof of one's own, consistent with a well-specified conceptual starting point. For a complete understanding of the coming sections, it is useful to recall the solution to which de Finetti himself always made reference, starting from the end of 1929. See [18] and [22].

He maintains probability assessments definitely have a subjectivistic source, but one can conceive the following ideal experiment to check on the closeness of a real-valued function P , defined on a class \mathcal{E} of events, to one's authentic opinions on the uncertainty of the elements of \mathcal{E} . One ought to think of $\{P(E) : E \in \mathcal{E}\}$ as a system of unit prices to have bets on the events included in \mathcal{E} . More precisely, if one is willing to accept bets of any amount (either positive or negative) at the unit prices fixed by the above system, then P represents the desired quantification of one's judgments. De Finetti assumes an ounce of rationality, in the sense that P is required to ensure that there is no choice of c_1, \dots, c_n in \mathbb{R} and of E_1, \dots, E_n in \mathcal{E} such that $\sup \sum_{k=1}^n c_k \{P(E_k) - \mathbb{1}_{E_k}\} < 0$, where \sup is taken with respect to all the elements (elementary events) of the partition of the sure event Ω generated by $\{E_1, \dots, E_n\}$, and $\mathbb{1}_E$ denotes the indicator function of the event E . Indeed, with this notation, $\sum_{k=1}^n c_k \{P(E_k) - \mathbb{1}_{E_k}\}$ represents the gain from a combination of bets of amount c_1, \dots, c_n on E_1, \dots, E_n , respectively. These remarks led de Finetti to restrict the class of admissible probability laws, on \mathcal{E} , to those which obey the *coherence principle*, i.e.:

For all *finite* subsets $\{E_i : i \in I\}$ of \mathcal{E} and $\{c_i : i \in I\}$ of \mathbb{R} , P satisfies

$$(1) \quad \sup \sum_{i \in I} c_i \{P(E_i) - \mathbb{1}_{E_i}\} \geq 0$$

with \sup taken with respect to all elementary events relative to $\{E_i : i \in I\}$.

Any P satisfying this property is said to be a *probability* on \mathcal{E} .

Existence of at least one probability law on \mathcal{E} is proved in [18]. It is of great interest the fact that coherence is well-defined, independently of the structure of \mathcal{E} . Moreover, as proved in [18], any probability on \mathcal{E} admits a coherent extension to any larger class of events.

As to the properties that one wishes to affirm, with a view to the calculus of probabilities, it is easy to *prove* that coherence entails:

- (π_1) If $\Omega \in \mathcal{E}$, then $P(\Omega) = 1$
- (π_2) $P(E) \geq 0$ for every E in \mathcal{E}
- (π_3) If E_1, E_2 and $E_1 \cup E_2$ are in \mathcal{E} , with $E_1 \cap E_2 = \emptyset$, then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

Moreover: If \mathcal{E} is an algebra, then (π_1), (π_2) and (π_3) are also sufficient in order that $P : \mathcal{E} \rightarrow \mathbb{R}$ can be considered a probability. See [22].

That is de Finetti's *proof* that countable additivity is not necessary for a function to be a probability on an algebra of events.

It would be misleading to consider countable families of bets with the purpose of extending additivity to countable families of pairwise disjoint events. Indeed, if one proceeds in this direction, it would be necessary to introduce extra-conditions, completely extraneous to the common interpretation of the term probability, in order to give a precise and unambiguous meaning to the gain.

To conclude, we mention a couple of facts that are of importance for understanding a few of the coming examples. Let $(P_n)_{n \geq 1}$ be a sequence of probabilities on \mathcal{E} and let $\mathcal{L} := \{E \in \mathcal{E} : \lim_{n \rightarrow \infty} P_n(E) \text{ exists}\} \neq \emptyset$. Then $Q(E) := \lim_{n \rightarrow \infty} P_n(E)$ is a probability on \mathcal{L} . It should be noted that Q is not necessarily countably additive even if each P_n is countably additive. The second fact to be considered herein is concerned with the general theory of stochastic processes. De Finetti has frequently referred to it, even if tacitly, in his already mentioned papers on random functions. Only at the end of the last century it was noticed and encapsulated into a theorem by Lester E. Dubins (1920–2010). See [28]. Define two (real-valued) random functions on $[0, +\infty)$ to be *cousins* if the family J of finite-dimensional p.d.'s of one of the functions is the same as the J of the other random function. Dubins proves that each random function, in particular the Brownian motion, has a cousin almost all of whose paths are polynomials, another cousin almost all of whose paths are step functions (on each bounded time-interval, they only have a finite number of values, each assumed on a finite union of intervals) that are continuous on the right (on the left), and a fourth cousin almost all of whose paths are continuous, piecewise-linear functions. Hence, in the next sections there will be no contradiction when, referring to de Finetti, continuous random functions will be considered with independent and stationary increments, different from the Brownian motion.

4. Presentation and critical comment of *Sui Passaggi*

The paper we are now going to analyze consists of nine sections. The first three are devoted to the explanation of some elementary facts concerning p.d. functions: these provide simple illustrations of how certain conclusions, valid for countably additive probabilities, do not generally hold any more for finitely additive probabilities. Sections 4 to 8 deal with convergence of sequences of random variables and include interesting, and somewhat amazing, remarks about the Cantelli strong law of large numbers. Finally, in Section 9, de Finetti goes back over the problem that had led him to the reflection developed in the previous sections: Is the p.d. of the integral of a random function equal to the limit of the laws of the integral sums?

4.1. Discussion of Sections 1–3

De Finetti starts with the p.d. function of a random number X , defined as

$$F(x) = P\{X < x\} \quad (x \in \mathbb{R})$$

and lingers on the correct interpretation of the discontinuity points of F , within a finitely additive framework. In order to avoid unnecessary complications, think of P as a probability on 2^Ω , Ω being a set and X a real-valued function defined on Ω . See the previous section for the definition and the existence of a probability on 2^Ω , and for de Finetti's attitude with respect to the general theory of probability at the end of 1929, when he was writing *Sui Passaggi*. In particular, he was perfectly aware of the existence of probabilities P such that $P(D) = 1$ and $P(\{x\}) = 0$ for some countable $D \subset \Omega$ and for every x in D . Since this will occur quite often in discussing de Finetti's paper, we now indicate a way to obtain probabilities of that type.

Example 4.1. Let D be a countably infinite subset of Ω , say $D = \{x_1, x_2, \dots\}$, and let P_n be a probability on 2^Ω defined by

$$P_n(A) := \frac{\sharp(A \cap D_n)}{n} \quad (A \subset \Omega)$$

with $D_n := \{x_1, \dots, x_n\}$. Write $\mathcal{L} := \{A \subset \Omega : \lim_{n \rightarrow \infty} \sharp(A \cap D_n)/n \text{ exists}\}$ and set $P(A) := \lim_{n \rightarrow \infty} P_n(A)$ for every A in \mathcal{L} . As recalled in the previous section, P is a probability on \mathcal{L} and, then, it admits a coherent extension γ on 2^Ω such that $\gamma(\{x_k\}) = 0$ and $\gamma(D) = 1$. \square

We are in a position to discuss the main issue of the present section, i.e. the correct interpretation of discontinuity for a p.d. function, within the frame established in Section 3. Under the ordinary condition of countable additivity, for any d of such type one would get $F(d-0) = F(d) = P\{X < d\} < F(d+0) = P\{X \leq d\}$ and, consequently, $F(d+0) - F(d-0)$ would represent the probability *concentrated* in $\{d\}$. This statement is not necessarily true if P is just finitely additive: in this case, one can only say that the following chain of inequalities holds true:

$$(2) \quad F(d-0) \leq P\{X < d\} \leq P\{X \leq d\} \leq F(d+0)$$

along with $F(d-0) < F(d+0)$ if d is a discontinuity point.

What de Finetti points out as a frequent mistake – in which, as recalled in Section 2, he himself had been trapped – the fact that many authors resorted to passages to the limit without preventive proof of the necessity of any condition justifying the limit processes along monotone sequences of events. Assuming either of the equalities $F(d-0) = P\{X < d\}$, $F(d+0) = P\{X \leq d\}$ might be an instance of that mistake, as displayed in the following example drawn from *Sui Passaggi*.

Example 4.2. Let $\Omega = \mathbb{R}$ and $(x_n)_{n \geq 1}$ be a sequence with $x_{n+1} < x_n$, for every n , and $x_n \downarrow 0$. Define D and γ as in the previous Example 4.1, and the random number X by $X(\omega) = \omega$ for every ω in $\Omega = \mathbb{R}$. Then, the p.d. function F of X satisfies $F(x) = 1 - P\{X \geq x\} = 1$ for every $x > 0$, and $F(x) = P\{X \leq x\} = 0$ for every $x \leq 0$. So, the jump ($= 1$) of F at $d = 0$ does not represent a mass concentrated in $\{0\}$. \square

This example shows that the inequality to the right of (2) cannot be replaced by equality, excepting special cases. To see that an analogous remark can be made for the inequality to the left, it is enough to consider the case of a sequence $(y_n)_{n \geq 1}$ with $y_n < y_{n+1}$ for every n and $y_n \uparrow 0$. Now, replacing P with a probability Q for which $Q(\{y_k\}) = 0$ for every k and $Q(\{y_1, y_2, \dots\}) = 1$, the consequent p.d. function G of X satisfies $G(x) = 0$ for $x < 0$ and $G(0) = 1 = P\{X < 0\}$. It should also be noted that $H := pF + (1-p)G$ is a p.d. function for each p in $[0, 1]$, and for every p in $(0, 1)$ inequalities under discussion hold on both sides of (2) when F is replaced by H and P by $pP + (1-p)Q$. It is also straightforward to find variants exhibiting concentrated masses, which satisfy

$$P\{X = d\} < F(d+0) - F(d-0).$$

As to the behaviour of a p.d. function F at $-\infty$ ($+\infty$, respectively), what can be said, in general terms, is that $\lim_{x \rightarrow -\infty} F(x) \geq 0$ ($\lim_{x \rightarrow +\infty} F(x) \leq 1$, respectively), strict inequalities being possible,

4.2. Discussion of Sections 4–9

In Sections 4–6, de Finetti discusses an important case in which, on the basis of the tacit assumption of continuity for probabilities, authors of the time were led to endow strong laws of large numbers with a meaning more general than one's due. De Finetti focuses on Cantelli's proof of the convergence of the frequency of success relative to n trials, in a sequence of Bernoulli trials, as n goes to infinity. To tackle the problem in the usual terms, define Ω to be the set of all sequences $d := (d_n)_{n \geq 1}$, each d_n being 0 or 1. Then, for each n define the n -th projection $p_n(d) := d_n$ ($d \in \Omega$), and set

$$f_n(d) = \frac{1}{n} \sum_{j=1}^n p_j(d).$$

Then, f_n represents the frequency of 1 in the first n trials. Let P be any probability on 2^Ω such that

$$P(\{d \in \Omega : p_1(d) = e_1, \dots, p_n(d) = e_n\}) = p^{\sum_{i=1}^n e_i} (1-p)^{n-\sum_{i=1}^n e_i}$$

where (e_1, \dots, e_n) is any point in $\{0, 1\}^n$, with $n = 1, 2, \dots$, p some fixed point in $[0, 1]$, and provided that $0^0 = 1$. Under these conditions, which imply that $(p_n)_{n \geq 1}$ is a Bernoulli sequence, Cantelli had proved that, for every $\varepsilon, \delta > 0$, there is $n_0 = n_0(\varepsilon, \delta)$ such that

$$(3) \quad \inf_{m \geq 1} P \left(\left\{ \max_{n \leq k \leq n+m} |f_k - p| \leq \varepsilon \right\} \right) \geq 1 - \delta$$

holds for any $n \geq n_0$.

It must be said that assessments of P for which

$$\inf_{m \geq 1} P \left(\left\{ \max_{n \leq k \leq n+m} |f_k - p| \leq \varepsilon \right\} \right) = P \left(\left\{ \max_{k \geq n} |f_k - p| \leq \varepsilon \right\} \right),$$

obtained by interchanging $\lim_{m \rightarrow \infty}$ with P , are consistent with the Bernoulli assumptions even in the frame of de Finetti's theory, but they are not the sole. In particular, there are assessments for which one cannot say that $(f_n)_{n \geq 1}$ converges to p with probability one, even though (3) is obviously valid. In Section 5 de Finetti illustrates the situation by means of the following interesting example.

Example 4.3. In the space Ω of all sequences d , each d_n being 0 or 1, define S_n to be the set of all sequences $(e_1, \dots, e_{n-1}, 1, 0, 0, \dots)$ obtained, for each $n \geq 2$, as (e_1, \dots, e_{n-1}) varies in $\{0, 1\}^{n-1}$. Moreover, set $S_1 := \{1, 0, 0, \dots\}$. So $\Omega_1 := \cup_{n \geq 1} S_n$ is the set of all sequences in Ω with a place occupied by 1 and followed by an entire sequence of 0's. Now, consider the sequence $(Q_n)_{n \geq 1}$ of probabilities on 2^Ω defined by $Q_1(S_1) = 1$ and, for any $n \geq 2$

$$Q_n(\{e_1, \dots, e_{n-1}, 1, 0, 0, \dots\}) = p^{\sum_{i=1}^{n-1} e_i} (1-p)^{n-1-\sum_{i=1}^{n-1} e_i}$$

for any (e_1, \dots, e_{n-1}) in $\{0, 1\}^{n-1}$. Clearly $Q_n(S_n) = 1$ for every $n \geq 1$. Finally, consider the probability γ of Example 4.1, with $\Omega = \mathbb{N} := \{1, 2, \dots\}$ and, for any subset A of Ω , set

$$Q(A) := \int_{\mathbb{N}} Q_n(A) \gamma(dn)$$

the integral being meant in the sense of Riemann-Stieltjes as explained, for example, in [47]. It is easy to verify that $Q_n(\{p_1 = e_1, \dots, p_N = e_N\}) = p^{\sum_{i=1}^N e_i} (1-p)^{N-\sum_{i=1}^N e_i}$ holds true for every $n \geq N+1$ and, since $\gamma(\{N+1, N+2, \dots\}) = 1$ for every N , one gets

$$\begin{aligned} Q(\{p_1 = e_1, \dots, p_N = e_N\}) &= \int_{\{n \geq N+1\}} Q_n(\{p_1 = e_1, \dots, p_N = e_N\}) \gamma(dn) \\ &= p^{\sum_{i=1}^N e_i} (1-p)^{N-\sum_{i=1}^N e_i}. \end{aligned}$$

This is tantamount to saying that $(p_n)_{n \geq 1}$ is a Bernoulli sequence with respect to Q . Then, (3) continues to be valid with Q in the place of P . Moreover, for each d in $\Omega_1 := \cup_{n \geq 1} S_n$ one has

$$f_n(d) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

Thus, since $Q(\Omega_1) = \int_{\mathbb{N}} Q_n(\cup_{k \geq 1} S_k) \gamma(dn) \geq \int_{\mathbb{N}} Q_n(S_n) \gamma(dn) = 1$, one sees that $(f_n)_{n \geq 1}$ converges to zero almost surely, and not to p . \square

The phenomenon highlighted in the previous example can be explained in the following terms. There is an instant n beyond which each trial turns into a failure ($= 0$). One is not able to predict when such an instant happens but, according

to the definition of γ , n is viewed as immensely distant. Hence, the probability of events which depend on any finite number of trials are not affected by the instant the sequence becomes a sequences of 0's. An analogous example has been provided, more recently, by Ramakrishnan and Sudderth in [46]. See also the instructive final *Remarks* therein, apropos of the common opinion on finitely additive probabilities.

As pointed out by de Finetti himself, one can change the sequence of 0's which, in Example 4.3, follows the last 1, in such a way that (3) still continues to be valid and, at the same time, f_n converges or does not: in the former case, it converges to a random variable with a prefixed p.d. function. For the sake of completeness, de Finetti's paper is here supplemented with a couple of additional examples of that type.

Example 4.4 (a) Let Ω be defined as in Example 4.3. Moreover, let $(y_n)_{n \geq 1}$ be the sequence defined by

$$y_j = \begin{cases} 0 & \text{if } j \in \{(2n-1)!, (2n-1)! + 1, \dots, (2n)! - 1\} \\ 1 & \text{if } j \in \{(2n)!, (2n)! + 1, \dots, (2n+1)! - 1\} \end{cases}$$

for $n = 1, 2, \dots$. Replace the sets S_1, S_2, \dots of Example 4.3 by

$$S_1^* := \{(1, y_1, y_2, \dots)\}$$

and, for $n \geq 2$,

$$S_n^* := \{(e_1, \dots, e_{n-1}, 1, y_1, y_2, \dots) : (e_1, \dots, e_{n-1}) \in \{0, 1\}^{n-1}\}.$$

Now, for each n , set

$$N_\nu := n + (2\nu)! - 1, \quad M_\nu := n + (2\nu + 1)! - 1 \quad (\nu = 1, 2, \dots).$$

Then, for every sequence in S_n^* , one has

$$(4) \quad \begin{cases} f_{N_\nu} = \frac{\{(2\nu-1)!+n-1\} f_{M_{\nu-1}}}{n+(2\nu)!-1} \rightarrow 0 & (\nu \rightarrow \infty) \\ f_{M_\nu} = \frac{\{(2\nu)!+n-1\} f_{N_\nu} + 2\nu(2\nu)!}{n+(2\nu+1)!-1} \rightarrow 1 & (\nu \rightarrow \infty). \end{cases}$$

Define probabilities Q_1^*, Q_2^*, \dots on 2^Ω according to

$$Q_1^*(\{(1, y_1, y_2, \dots)\}) = 1$$

and, for $n \geq 2$,

$$Q_n^*(\{e_1, \dots, e_{n-1}, 1, y_1, y_2, \dots\}) = p^{\sum_{i=1}^{n-1} e_i} (1-p)^{n-1-\sum_{i=1}^{n-1} e_i}.$$

Then, set

$$Q^*(A) := \int_{\mathbb{N}} Q_n^*(A) \gamma(dn) \quad (A \subset \Omega).$$

It is easy to prove that $(p_n)_{n \geq 1}$ is a Bernoulli sequence even with respect to Q^* . Hence, (3) is valid even with Q^* in the place of Q . On the other hand, in view of (4), for each sequence in Ω_1^* one has $\liminf f_n = 0 < \limsup f_n = 1$, and $Q(\Omega_1^*) = 1$.

(b) Maintain the meaning for Ω , γ and $(f_n)_{n \geq 1}$. Denote by C the subset of Ω on which $(f_n)_{n \geq 1}$ converges and by H a prefixed p.d. function, supported by $[0, 1]$. Finally, let σ be a σ -additive probability measure on $\mathcal{B}(\{0, 1\}^\infty \cap C)$ such that

$$\sigma(\{p_1 = e_1, \dots, p_n = e_n\}) = \int_{[0,1]} \theta^{\sum_{i=1}^n e_i} (1 - \theta)^{n - \sum_{i=1}^n e_i} dH(\theta)$$

for any $(e_1, \dots, e_n) \in \{0, 1\}^n$ and $n \geq 1$. Now, define

$$\begin{aligned} S_n^{**} &:= \{(e_1, \dots, e_{n-1}, 1, s) : (e_1, \dots, e_{n-1}) \in \{0, 1\}^{n-1}, s \in C\} \quad (n \geq 2) \\ S_1^{**} &:= \{(1, s) : s \in C\} \end{aligned}$$

together with the probabilities

$$\begin{aligned} Q_1^{**}(\{1\} \times A) &:= \sigma(A) \\ Q_n^{**}(\{e_1, \dots, e_{n-1}, 1\} \times A) &= \beta_{n-1}(\{e_1, \dots, e_{n-1}\}) \sigma(A) \end{aligned}$$

for any A in $\mathcal{B}(\{0, 1\}^\infty \cap C)$, $(e_1, \dots, e_{n-1}) \in \{0, 1\}^{n-1}$, $n \geq 2$ and

$$\beta_{n-1}(\{e_1, \dots, e_{n-1}\}) = p^{\sum_{i=1}^{n-1} e_i} (1 - p)^{n-1 - \sum_{i=1}^{n-1} e_i}.$$

Finally,

$$Q^{**}(A) := \int_{\mathbb{N}} Q_n^{**}(A) \gamma(dn) \quad (A \subset \Omega).$$

Once again, $(p_n)_{n \geq 1}$ turns out to be a Bernoulli sequence with respect to Q^{**} and, then, (3) is valid with Q^{**} in the place of Q . On the other hand, in view of de Finetti's theory of exchangeable sequences, $Q^{**}(\Omega_1^{**}) = 1$ with $\Omega_1^{**} = \cup_{n \geq 1} S_n^{**}$, and $(f_n)_{n \geq 1}$ converges, on Ω_1^{**} , to a random number whose p.d. function is H . \square

The last two groups of examples have an important element in common. They show that there are sequences of random numbers either converging almost surely or oscillating almost surely, with p.d. functions converging weakly in both cases to a p.d. function which differs from the p.d. function of the almost sure limit: The former, in those examples, has a jump invariably equal to 1 at p , whereas the latter can be let varying in the class of all p.d. functions. De Finetti explains why he was interested in investigating into these phenomena in the last section of *Sui Passaggi*. In the second half of 1929 he was about to deduce the p.d. of the integral of a continuous stochastic f.i.s.i. as limit of the p.d. functions of integral sums which converge pointwise to the integral of interest. At this point, he was assailed by the doubt that such a line of reasoning could be in conflict with his way of thinking of probability, in the sense that the argument could be valid for specific extensions of a prefixed system of finite-dimensional laws, but not in general. The above examples

confirmed the reasonableness of his doubt. Here we faithfully follow [13] and provide the reader with some further insight on this aspect.

Let $t \mapsto X(t)$ be a random function, for $t \geq 0$, with continuous trajectories such that $X(0) \equiv c$. See Section 3. Then, for each $t > 0$, one can write

$$\int_0^t X(u) du = \lim_{n \rightarrow \infty} \frac{t}{n} \sum_{h=1}^n X\left(\frac{th}{n}\right)$$

where, by the Brunacci–Abel identity,

$$V_n := \frac{t}{n} \sum_{h=1}^n X\left(\frac{th}{n}\right) = ct + \frac{t}{n} \sum_{h=1}^n (n-h+1) \left\{ X\left(\frac{th}{n}\right) - X\left(\frac{t(h-1)}{n}\right) \right\}.$$

Then, if X is a f.i.s.i., as proved in [11], the characteristic function of the increment $\{X(\frac{th}{n}) - X(\frac{t(h-1)}{n})\}$ is given by $\phi(\cdot)^{t/n}$, with $\phi(\cdot)$ being the characteristic function of $X(1)$. Then, for the characteristic function Ψ_{V_n} of V_n one gets

$$\text{Log } \Psi_{V_n}(\xi) = ic\xi t + \frac{t}{n} \sum_{h=1}^n \text{Log } \phi\left((n-h+1)\frac{t}{n}\xi\right) \quad (\xi \in \mathbb{R})$$

where Log denotes the principal branch of the logarithm. Then, $\lim_{n \rightarrow \infty} \Psi_{V_n}$ exists, uniformly on compact intervals, and is given by

$$(5) \quad \exp \left\{ ic\xi t + \frac{1}{\xi} \int_0^{\xi t} \text{Log}(\phi(u)) du \right\}.$$

After proving that such a limit is a characteristic function, de Finetti can assert that V_n converges in distribution, but he cannot state that (5) is the characteristic function of $\lim_n V_n = \int_0^t X(u) du$. In fact, in [2] it is shown that, in the frame of de Finetti's theory as summarized in Section 3, $\int_0^1 X(u) du$ can be given any p.d. when, for example, X has the finite-dimensional distributions of the standard Brownian motion.

As a natural development of the previous remarks, de Finetti tries to find interesting types of convergence of a sequence $(X_n)_{n \geq 1}$ of random numbers to a random number X , which entails weak convergence of the p.d. functions of the X_n 's to the p.d. of X . It is apropos of this question that he introduced the general notion of stochastic uniform convergence (namely convergence in probability) in [13] and proved that it meets the above property, with respect to weak convergence, in Section 8 of [14]. See the previous Section 2. Coming back to the problem of the law of the integral of a random function, he concludes *Sui Passaggi* touching on special cases in which one can verify that $(V_n)_{n \geq 1}$ converges in probability to $\int_0^t X(u) du$. This is the case when, for example, X is non-decreasing.

5. Ensuing correspondence between de Finetti and Fréchet

The correspondence consists of four short open letters published in *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* and presented in two meetings of the

Istituto: The first and the second, by Fréchet and de Finetti respectively, were presented on the 3rd of July 1930, whereas the third and the fourth, by Fréchet and de Finetti respectively, on the 20th of November 1930.

In his first note, Fréchet begins by saying that *Sui Passaggi* is an interesting paper, which has points of contact with topics he has been dealing with in his recent courses (1929-1930). These were summarized in an article still in press at the beginning of July: that is the paper [31], mentioned in Section 2 and published in last issue of *Metron* of that very same year. Fréchet agrees with de Finetti on the fact that countable additivity cannot be deduced from finite additivity, and that the latter constitutes a principle generally accepted as a basis for the theory and the calculus of probabilities. But he has a different opinion about the admissibility of probabilities that are not countably additive. He explains this attitude by referring to the alternative arisen to the founders of the modern theory of measure. They, in spite of the awareness of the impossibility of the problem of measure highlighted by the celebrated Vitali example (see [52]), opted for countable additivity restricted to suitable domains. Fréchet holds this way of proceeding up as an example, and proposes introducing the idea of “événements qui ont une probabilité déterminée et d’autres qui n’en auront pas”, provided that the condition of countable additivity is admitted by definition. Then, he asserts that in such a case, the events considered by de Finetti in his examples do not have a “probabilité déterminée”, and concludes that it is in this circumstance that the solution to the questions raised by de Finetti must be sought.

In his answer, de Finetti respectfully tries to bring the debate down to the real question: To decide if all finitely additive probabilities are admissible, or alternatively if it is necessary to restrict admissible probabilities to the laws that are countably additive. He says he has the sense that several authors, dealing with this subject, consider themselves free to decide according to what suits them best. As an example of this attitude, he mentions the Fréchet evocation of an analogy with the theory of measure. Taking his cue from this, de Finetti says he considers it unjustified to make use of *conventions* to define concepts, like probability, that have a proper meaning, even if possibly open to dispute. Then, the main issue does not consist of making more or less arbitrary conventions on certain properties, but can be traced back to *proving* that certain properties are *necessary*. As mentioned in Section 3, de Finetti derived the necessity of finite additivity from a coherence principle which, far from being merely conventional, corresponds to a prevailing rational attitude. To say that an event E has probability p either has a more or less common intuitive meaning, or is a perfectly useless sentence. In the first case, if we have a countable family of mutually disjoint events (E_1, E_2, \dots) , with $P(E_n) = 0$ for every n , are we able to conclude that $E := \cup_{n \geq 1} E_n$ has probability 0, or, equivalently, that E is invariably practically impossible? It is plain that this is not a convention matter, since the above conclusion has a real conceptual content. It can be false or true, but this must be proved, not assumed by a convention. De Finetti admits a

quite serious difficulty in doing this, due to the lack, until then, of a general theory of probability. He takes the advantage of this circumstance to announce that he has completed a theory, which, starting from a method for assessing probabilities [betting scheme, auth. note], allows one to impeccably deduce the mathematical properties of probability. This is clearly a reference to the theory outlined in the previous Section 3. He recalls that, within such a theory, countable additivity is not a necessary requisite for admissibility and, then, all the examples given in *Sui Passaggi* are perfectly justified and make sense. He concludes with a mention to a couple of consequences, that he considers bizarre, of the adoption of countable additivity as a compulsory principle for admissibility of probability laws. The first consequence relates to the fact that such a principle would forbid one to think of a sequence of mutually disjoint events, forming a partition of the sure event and having probabilities of the same order of magnitude. In other words, that sequence invariably ought to include a finite subset with respect to which the whole of the remaining events would be negligible. The second bizarre consequence of assuming, as a compulsory principle, countable additivity is that one cannot say that the weak limit of a sequence of p.d. functions is always a p.d. function.

The “official” reply by Fréchet to the previous de Finetti’s arguments was submitted for publication on the 20th of November 1930. The abstract is categoric but the content of the paper is kept to the point more than the first letter. As to the abstract, he writes that de Finetti’s examples are inadmissible, on the basis of the following facts: First, the probabilities of the events considered therein cannot be expressed by real numbers. Second, the probability laws studied therein are inconsistent with empirical experience. Fréchet splits his criticism into four points, and de Finetti answers them in the same order. The final part of this section is accordingly organized into four subsections, each of which summarizes both the Fréchet critical remark and the de Finetti answer pertaining to the point in title of the subsection.

Point 1. Fréchet (F in the sequel) says that de Finetti (dF in the sequel) has proved, by simple examples, that it is not possible to define a probability for *all* the events while complying, at the same time, with the condition of countable additivity for all classes of mutually disjoint events. Then, he insists on the point, already mentioned in his first letter, that the question can be solved by introducing suitable restrictions on the class of the events equipped with a probability, in such a way that countable additivity is preserved.

dF replies that F has misunderstood his thought. Indeed, he maintains, on the one hand, that it is always possible to comply with the principle of countable additivity and that, on the other hand, all the laws that meet the principle of coherence are admissible. The first part of this argument is not in contrast with the Vitali theorem, which simply excludes that a countably additive law might give equal probabilities to all superposable subsets of $[0, 1]$. Then, to overcome the drawback, it is not necessary to restrict the class of the admissible events. Moreover,

such a strategy is not so much as sufficient. To see this, reconsider Example 4.1 with $\Omega = (0, 1]$ and $x_1 = 1, x_2 = 1/2, x_3 = 1/3, \dots$ and form the partition of $(0, 1]$ in the intervals given by $(1/2, 1], (1/3, 1/2], \dots, (1/(n+1), 1/n], \dots$. Then, since $P(\{x_1, \dots, x_n, \dots\}) = 1$, one gets $P((0, 1]) = 1$, but $P((1/(n+1), 1/n]) = 0$ for every n . In other words, the “drawback” of countable partitions of Ω , into elements having zero probabilities, can occur even if these elements are intervals, which, according to F, constitute the most typical example of “probabilizable events”. It is at this point that dF drops the artificial examples suggested by the theory of measure, to explain how more concrete situations, pertaining to the study of functions with random increments, are open to the same objections as those raised in the previous more schematic examples. To this end, dF mentions two cases. One of them involves the notion of *conglomerability* and, hence, goes beyond the scope of the present paper. The original paper in which that phenomenon had been noticed is [20]. The other case is concerned with the already mentioned de Finetti’s theorem on the nowhere differentiability of the trajectories of the Brownian motion. See Section 2. He expresses his regret that, in view of his criticism on the role of countable additivity, the aforementioned theorem must be reformulated in a weaker form, i.e.: Let ε and M be strictly positive numbers. Then, the probability that $[0, 1]$ includes any interval of length greater than ε , for which one gets $|X(t_2) - X(t_1)| < M(t_2 - t_1)$, is zero. Since the usual formulation, obtained as $M \rightarrow \infty$ and $\varepsilon \rightarrow 0$, would be very important, if valid in general, dF admits he would be very happy of the existence of any reasonable argument that persuades him to share the common idea that probability laws are continuous. He concludes expressing his skepticism towards the solution devised by F: “Even supposing that there are events for which the doubts about countable additivity turn out to be groundless, how could I recognize them in practical situations of the same type as that just now described.”

Point 2. This corresponds to the first point raised in the above-mentioned abstract. F seems to admit that restricting the class of the “probabilizable events” does not serve the purpose to explain the antinomy stressed by de Finetti’s examples. In order to solve the issue, he proposes to consider probabilities expressed in terms of actual infinitesimals, say ε . He supposes that, in such a way, one succeeds in writing $\varepsilon \cdot \omega = 1$. dF observes that the adoption of “new numbers” of the type of ε is not inconsistent with his own way of thinking. In fact, he recalls he has already made use of those numbers in [10]. But, unlike F, he gets to the conclusion there is no contradiction between admissibility of infinitesimals and finite additivity. Indeed, the probability of the union of any finite number of infinitesimal events is infinitesimal and no limit process could lead to conclude that the probability of the union of all the events is 1.

Point 3. It corresponds to the second point briefly described in the abstract. F maintains that probabilities like those of Example 4.1 do not appear when prob-

abilities are based on frequencies. Moreover, he is skeptical about the success of dF's scientific plan, since none of the definitions of probability proposed until then had met with general approval. So, he refers the reader to the final part of the Lévy monograph [41], where countable additivity is, according to his opinion, justified. Apropos of the first assertion, he claims he is able to prove it in the following terms. Consider a random phenomenon with a countable set of elementary possible outcomes, say a_1, a_2, \dots , and assume one can conceive an indefinitely extendable sequence of trials of that phenomenon. Let $f_k^{(n)} = r_k/n$ be the frequency of a_k in the first n trials. Then, all but a finite number of the $f_k^{(n)}$'s are zero and $f_1^{(n)} + f_2^{(n)} + \dots = 1$. It is palpably clear that such an equality will be valid for any n . According to the empirical interpretation, as explained in [33], $f_k^{(n)}$ represents an "experimental measure" of a probability p_k , for every k , as n increases. F claims that these remarks are sufficient to conclude that $p_1 + p_2 + \dots = 1$.

In his reply, dF stresses once again that properties like countable additivity be proved and not established in the form of conventions. Even if he admits the existence of the difficulties mentioned by F, he considers them more extrinsic rather than intrinsic. He says that, to clear this hurdle, it would suffice that every author would give his own proof, based on his own definition of probability. He acknowledges that this is what F has tried to do, starting from his empirical interpretation of probability. But, despite the inadequacy of the F proof, dF shows that the F argument, once made more precise, can become an excellent point in favour of the dF thesis. As to the inadequacy, assuming that each p_k is the (usual) limit, as $n \rightarrow \infty$, of $(f_k^{(n)})_{n \geq 1}$, one can write $1 = \lim_{n \rightarrow \infty} \sum_k f_k^{(n)}$ but, in general, the exchange of \lim with \sum is not valid. Hence, one can just say, in general, that $1 \geq \sum_k p_k$. In the footnote (2) of page 256, dF notes that F had pointed out, in personal correspondence, that he didn't mean to speak of p_k as a limit in a mathematical sense. Once taken note of this detail, dF resorts to a different argument free from the criticism of being just in an abstract mathematical form. He sets himself the objective of studying the expected behaviour of the $f_k^{(n)}$'s, as n goes to infinity, to show that *there are probability laws with respect to which it would be illusory to expect that the $f_k^{(n)}$'s converge to numbers p_k such that $\sum p_k = 1$* . Therefore, in the same way, it would be illusory to hope to prove that the property of countable additivity may be derived, in general, from the analogous property valid for frequencies, understood as *empirical* estimates of probabilities. Here is the example proposed by dF along with a few further details.

Example 5.1. Let S be a countably infinite set, say $S := \{a_1, \dots, a_n, \dots\}$, and $\Omega = S^\infty$. Define ξ_1, ξ_2, \dots to be the coordinate random variables of Ω . Now, for any strictly positive integer N , set $S_N := \{a_1, \dots, a_N\}$ and, for every $A \subset S^n$, define

$$P_N^{(n)}(A) = \frac{\#(A \cap S_N^n)}{N^n}.$$

It is plain that $(S^n, 2^{S^n}, P_N^{(n)})_{n \geq 1}$ is a consistent system of probabilities. Then, there exists a probability P_N on $(\Omega, 2^\Omega)$ such that $P_N(A \times S^\infty) = P_N^{(n)}(A)$ holds

true for every $A \subset S^n$, $n = 1, 2, \dots$. Finally, with the same γ as in Example 4.3, put

$$P(C) = \int_{\mathbb{N}} P_N(C) \gamma(dN) \quad (C \subset \Omega).$$

Hence, P is a probability on 2^Ω such that $P\{\xi_n = a_k\} = 0$ for every k and n . Indeed, for every $N \geq k$, one has

$$P_N\{\xi_n = a_k\} = P_N^{(n)}(S^{n-1} \times \{a_k\}) = \frac{\#((S^{n-1} \times \{a_k\}) \cap S_N^n)}{N^n} = \frac{N^{n-1}}{N^n}$$

and, then, $P\{\xi_n = a_k\} = \int_{\{N \geq k\}} N^{-1} \gamma(dN) = 0$. Moreover, for any $n \geq 2$,

$$(6) \quad P\{\xi_1 \neq \xi_2 \neq \dots \neq \xi_n\} = 1$$

Indeed, if $N > n$, one gets

$$\begin{aligned} P_N(\{\xi_1 \neq \xi_2 \neq \dots \neq \xi_n\} \times S^\infty) &= P_N^{(n)}\{\xi_1 \neq \xi_2 \neq \dots \neq \xi_n\} \\ &= \frac{N(N-1) \dots (N-n+1)}{N^n} \end{aligned}$$

and, since

$$P\{\xi_1 \neq \xi_2 \neq \dots \neq \xi_n\} = \int_{\{N \geq n\}} \frac{N(N-1) \dots (N-n+1)}{N^n} \gamma(dN)$$

the equality in (6) follows. Now, if $(p_k)_{k \geq 1}$ is any sequence such that $p_k \geq 0$ for every k and $\sum_k p_k = 1$, there is \bar{k} such that $p_{\bar{k}} = \max\{p_1, p_2, \dots\} > 0$. Then, there is a contradiction between the adoption of P as a probability on 2^Ω and the assumption that $f_k^{(n)} := \sum_{j=1}^n \mathbb{1}_{\{a_k\}}(\xi_j)/n$ converges, in some sense, to p_k for every k , as $n \rightarrow \infty$. Indeed, with respect to P , for each n it is practically sure – in view of (6) – that there are n indices, say k_1, \dots, k_n , for which $f_{k_i}^{(n)} = 1/n$ ($i = 1, \dots, n$). So, one gets

$$(7) \quad P\left\{f_{\bar{k}}^{(n)} \leq \frac{p_{\bar{k}}}{M}\right\} = 1$$

for any $M, n \in \mathbb{N}$ and $n > M/p_{\bar{k}}$, thus contradicting any reasonable definition of convergence of $f_{\bar{k}}^{(n)}$ to $p_{\bar{k}}$, as $n \rightarrow \infty$. \square

The above example shows that the condition of countable additivity does not follow from the fact that $\sum_k f_k^{(n)} = 1$ holds true in any case, combined with the assumption that frequencies approach probabilities as n increases to infinity. Indeed, (7) and the arbitrariness of $(p_k)_{k \geq 1}$ show, once again, the inadequacy of the use of frequencies to prove countable additivity. The sense of this statement can be strengthened even further when P_N is defined to be the Kolmogorov extension of the $P_N^{(n)}$'s to the smallest σ -algebra containing all the sets $\xi_m^{-1}(A)$ for all m and all $A \subset S$. It is easy to check that the ξ_n 's turn out to be independent and identically distributed, with uniform distribution on S_N , with respect to any extension of the

$P_N^{(n)}$'s. Moreover, if P_N corresponds to the Kolmogorov extension, then the strong law of large numbers yields

$$P_N \left\{ \lim_n f_k^{(n)} = \frac{1}{N} \text{ for } k = 1, \dots, N, \quad \lim_n f_k^{(n)} = 0 \text{ for } k \geq N+1 \right\} = 1$$

and, with the same $(p_n)_{n \geq 1}$ as in Example 5.1 and for every M in \mathbb{N} one obtains

$$\begin{aligned} P \left\{ \sum_k \overline{\lim}_{n \rightarrow \infty} f_k^{(n)} \leq \frac{1}{M} \right\} &\geq P \left\{ \overline{\lim}_{n \rightarrow \infty} f_k^{(n)} \leq \frac{p_k}{M} \text{ for every } k \right\} \\ &= \int_{\mathbb{N}} P_N \left\{ \overline{\lim}_{n \rightarrow \infty} f_k^{(n)} \leq \frac{p_k}{M} \text{ for every } k \right\} \gamma(dN) \\ &= 1. \end{aligned}$$

Then, with respect to this particular P , we are practically sure that frequencies converge and that the sum of the series is in $[0, \varepsilon)$ for every $\varepsilon > 0$. \square

Point 4. The last objection raised by F is about a seeming slip made by dF, in the previous part of the correspondence, apropos of the nature of the weak limit of a sequence of p.d. functions. As already recalled in this very same section, dF found it bizarre that, in the common approach based on countable additivity, such limit was not necessarily a p.d. function. F maintains that this phenomenon is not warded off by the adoption of dF point of view. In support of this statement, F gives the example of the p.d. functions $F_n(x) = (\mathbf{1}_{(-n, n]}/2 + \mathbf{1}_{(n, \infty)})(x)$, $x \in \mathbb{R}$ and $n = 1, 2, \dots$. It is easy for dF to prove that the F argument is ineffective since the function $F \equiv 1/2$, limit of $(F_n)_{n \geq 1}$, can be viewed, within the finitely additive frame, as a p.d. function.

6. Final remarks

The paper we have annotated in the previous sections cannot be counted among de Finetti's most important works. Nonetheless, it represents a direct evidence of an extremely interesting stage of his scientific career. It is the stage of the mathematical formulation of his subjectivistic conception of probability, and of the consequent conclusion that the only general restriction on the class of the admissible probability laws is given by the coherence principle. Hence, probabilities on algebras of events must be additive, but not necessarily countably additive. This, on the one hand, led him to revise the value, in terms of their generality, of a few of his own previous theorems that were proved, in part, under the assumption of continuity of probability laws. His critique, on the other hand, did not even spare one of the most renowned achievements of the theory of probability, i.e. the strong law of large numbers. By resorting to enlightening examples, in *Sui Passaggi* he succeeds in enhancing some crucial differences between the two viewpoints taken into consideration therein. Besides, the discussion with Fréchet gives de Finetti a chance to provide fresh and deep explanations about his stance. They continue to be of

great interest and useful since the arguments, still put up against the adoption of de Finetti's theory, do not basically differ from those used by Fréchet.

References

- [1] BASSETTI, F. and REGAZZINI, E. (2008). The unsung de Finetti's first paper about exchangeability. *Rendiconti di Matematica* **28**, 1–17.
- [2] BERTI, P., REGAZZINI, R. and RIGO, P. (2007). Modes of convergence in the coherent framework. *Sankhyā* **69**, 314–329.
- [3] CANTELLI, F.P. (1916). La tendenza a un limite nel senso del calcolo delle probabilità. *Rendiconti del Circolo Matematico di Palermo* **41**, 191–201.
- [4] CANTELLI, F.P. (1917). Sulla probabilità come limite della frequenza. *Atti R. Accademia Nazionale dei Lincei, Serie V, Rend.* **26**.
- [5] CANTELLI, F.P. (1935a). Considérations sur la convergence dans le calcul des probabilités. *Annales de l'Institut Henri Poincaré* **5**, 3–50.
- [6] CANTELLI, F.P. (1935b). Sulla estensione del principio delle probabilità totali a una successione illimitata di eventi incompatibili. *Giornale Istituto Italiano degli Attuari* **6**, 415–427.
- [7] CANTELLI, F.P. (1936). Sulla estensione del principio delle probabilità totali a una successione illimitata di eventi incompatibili. In *Scritti Matematici offerti a Luigi Berzolari* (Brusotti, L., Palatini, A., Serini, R., Gallina, G., eds.). *Istituto Matematico della R. Università di Pavia*, Pavia.
- [8] CASTELNUOVO, G. (1919). *Calcolo delle Probabilità e Applicazioni.*, Soc. Ed. Dante Alighieri, Milano.
- [9] CIFARELLI, D.M. and REGAZZINI, E. (1996). De Finetti's contribution to probability and statistics. *Statist. Sci.* **11**, 253–282.
- [10] DE FINETTI, B. (1928). Sulle probabilità numerabili e geometriche. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **61**, 817–824.
- [11] DE FINETTI, B. (1929a). Sulle funzioni a incremento aleatorio. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **10**, 163–168.
- [12] DE FINETTI, B. (1929b). Sulla possibilità di valori eccezionali per una legge di incrementi aleatori. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **10**, 325–329.
- [13] DE FINETTI, B. (1929c). Integrazione delle funzioni a incremento aleatorio. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **10**, 548–553.
- [14] DE FINETTI, B. (1930a). Sui passaggi al limite nel calcolo delle probabilità. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **63**, 155–166.
- [15] DE FINETTI, B. (1930b). A proposito dell'estensione del teorema delle probabilità totali alle classi numerabili. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **63**, 901–905.
- [16] DE FINETTI, B. (1930c). Ancora sull'estensione alle classi numerabili del teo-

- rema delle probabilità totali. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **63**, 1063–1069.
- [17] DE FINETTI, B. (1930d). Le funzioni caratteristiche di legge istantanea. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **12**, 278–282.
 - [18] DE FINETTI, B. (1930e). Problemi determinati e indeterminati nel calcolo delle probabilità. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **12**, 367–373.
 - [19] DE FINETTI, B. (1930f). Funzione caratteristica di un fenomeno aleatorio. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Mem.* **4**, 86–133.
 - [20] DE FINETTI, B. (1930g). Sulla proprietà conglomerativa delle probabilità subordinate. *Atti R. Istituto Lombardo di Scienze e Lettere* **63**, 414–418.
 - [21] DE FINETTI, B. (1931a). Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza. *Logos* (Biblioteca di Filosofia, diretta da A. Aliviera), 163–219- Pezzella, Napoli.
 - [22] DE FINETTI, B. (1931b). Sul significato soggettivo della probabilità. *Fundamenta Mathematicae* **17**, 298–329.
 - [23] DE FINETTI, B. (1931c). Sul concetto di media. *Giornale Istituto Italiano Attuari* **2**, 369–396.
 - [24] DE FINETTI, B. (1937). A proposito di “correlazione”. *Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **3**, 41–57.
 - [25] DE FINETTI, B. (1939). La teoria del rischio e il problema della “rovina dei giocatori”. *Giornale Istituto Italiano Attuari* **10**, 41–51.
 - [26] DE FINETTI, B. (1981). *Scritti*. Cedam, Padova.
 - [27] DELL’AGNOLA, C.A. (1929). Intorno alle successioni di variabili casuali discontinue tendenti ad una variabile casuale limite. *Atti R. Istituto Veneto Scienze e Lettere* **88**, 40–62.
 - [28] DUBINS, L.E. (1999). Paths of finitely additive Brownian motion need not be bizarre. *Séminaire de Probabilités XXXIII* (Azéma, J., Emery, M., Ledoux, M., Yor, M., eds.) Lecture Notes in Mathematics 1709. Springer, Berlin.
 - [29] FRÉCHET, M. (1930a). Sur l’extension du théorème des probabilités totales au cas d’une suite infinie d’événements. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **63**, 899–900.
 - [30] FRÉCHET, M. (1930b). Sur l’extension du théorème des probabilités totales au cas d’une suite infinie d’événements. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* **63**, 1059–1062.
 - [31] FRÉCHET, M. (1930c). Sur la convergence en probabilité. *Metron* **8**, 1–50.
 - [32] FRÉCHET, M. (1937). *Généralités sur les probabilités. Variables Aléatoires*. Gauthier-Villars, Paris.
 - [33] FRÉCHET, M. and HALBWACHS, M. (1924). *Le Calcul des Probabilités à la portée de tous*. Dunod, Paris.
 - [34] KHINCHIN, A. (1932a). Sur les classes d’événements équivalents. *Mat. Sb.* **39**, 73–84.

- [35] KHINCHIN, A. (1932b). Remarques sur les suites d'événements obéissants à la loi des grands nombres. *Mat. Sb.* **39**, 115–119.
- [36] KHINCHIN, A. (1937). A new derivation of a formula of P. Lévy. *Bull. Moscow Gov. Univ.* **1**, 1–5.
- [37] KOLMOGOROV, A.N. (1930). Sur la notion de la moyenne. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **12**, 388–391.
- [38] KOLMOGOROV, A.N. (1932a). Sulla forma generale di un processo stocastico omogeneo (un problema di Bruno de Finetti). *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **15**, 805–808.
- [39] KOLMOGOROV, A.N. (1932b). Ancora sulla forma generale di un processo omogeneo. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend.* **15**, 866–869.
- [40] KOLMOGOROV, A.N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Ergebnisse der Mathematik und ihrer Grenzgebiete **3**. Springer, Berlin.
- [41] LÉVY, P. (1925). *Calcul des Probabilités*. Gauthier-Villars, Paris.
- [42] LÉVY, P. (1934). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Sup. Pisa* **3**, 337–366.
- [43] LÉVY, P. (1970). *Quelques aspects de la pensée d'un mathématicien*. Blanchard, Paris.
- [44] OTTAVIANI, G. (1939). Sulla teoria astratta del calcolo delle probabilità proposta dal Cantelli. *Giornale Istituto Italiano degli Attuari* **10**, 133–163.
- [45] PALEY, R.E.A.C., WIENER, N. and ZYGMUND, A. (1933). Notes on random functions. *Math. Z.* **37**, 647–668,
- [46] RAMAKRISHNAN, S. and SUDDERTH, W.D. (1988). A sequence of coin toss variables for which the strong law fails. *Amer. Math. Monthly* **95**, 939–941.
- [47] RAO, K.P.S. BHASKARA and RAO, M. BHASKARA (1983). *Theory of Charges*. Academic Press, London.
- [48] REGAZZINI, E. (2005). Probability and statistics in Italy during the First World War. I: Cantelli and the laws of large numbers. *Journal Electronique d'Histoire des Probabilités et de la Statistique* **1**, 1–12.
- [49] REGAZZINI, E. (2007). De Finetti's contribution to the theory of random functions. *International Journal of Approximate Reasoning* **47**, 387–393.
- [50] SENETA, E. (1992). On the history of the strong Law of Large Numbers and Boole's inequality. *Historia Mathematica* **19**, 24–39.
- [51] SLUTSKY, E. (1925). Ueber stochastische Asymptoten und Grenzwerte. *Metron* **5**, 1–90.
- [52] VITALI, G. (1905). Sul problema della misura dei gruppi di punti di una retta. Tipografia Gamberini e Parmeggiani, Bologna.
- [53] VON MISES, R. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Springer, Wien.